

YILUN DU

Harvard Science and Engineering Complex
150 Western Ave, 6.138
Allston, MA 02134

ydu@seas.harvard.edu
<https://yilundu.github.io/>

Research Overview

My research seeks to build intelligent agents that can understand, reason about, and act effectively in complex, unfamiliar physical environments. Because available decision-making data cannot span the full range of situations an agent may encounter, I study how agents can learn reusable models, then compose and reason with them at inference time. Composition allows the same learned knowledge to support new combinations of world dynamics, goals, and concepts, enabling agents to solve tasks not encountered during training. I develop generative models that represent these components, together with inference methods for searching, optimizing, and planning over their compositions. My work has demonstrated the value of this approach across perception, reasoning, and robot planning, with broader applications across science and engineering.

Research Areas: Generative Modeling, Inference-Time Reasoning, Compositional Learning, World Models, Robot Learning, and Multi-Agent Intelligence.

Education

Ph.D. EECS 2020-2024	Massachusetts Institute of Technology (MIT) Advisors: Leslie Kaelbling, Tomás Lozano-Pérez, and Joshua B. Tenenbaum Thesis: <i>Learning Generalizable Systems by Learning Composable Energy Landscapes</i>
M.S. EECS 2019-2020	Massachusetts Institute of Technology (MIT) Advisors: Leslie Kaelbling, Tomás Lozano-Pérez, and Joshua B. Tenenbaum Thesis: <i>Implicit Learning with Energy-Based Models</i>
B.S. 2015-2019	Massachusetts Institute of Technology (MIT) EECS, GPA: 5.0 / 5.0

Appointments

Assistant Professor of Computer Science July 2025–Present	Harvard University Harvard John A. Paulson School of Engineering and Applied Sciences Institute Investigator, Kempner Institute Director, Embodied Minds Lab
Part-Time Research Scientist Spring 2026–Present	NVIDIA Research Learning and Perception Research Group
Senior Research Scientist 2024–2025	Google DeepMind Robotics Team

Prior Research Experience

Graduate Researcher 2019–2024	Massachusetts Institute of Technology Learning and Intelligent Systems Group, CSAIL Computational Cognitive Science Lab, Brain and Cognitive Sciences Advisors: Tomás Lozano-Pérez, Leslie Pack Kaelbling, and Joshua B. Tenenbaum
Research Intern Summer 2023	Google Robotics Video models for robotics
Student Researcher 2022–2023	Google Brain Video models for decision making
Research Intern Summer 2022	DeepMind Compositional generative modeling
Visiting Researcher 2020–2022	Facebook AI Research FAIR Proteins (Generative Biology) Team
Research Intern Summer 2019	Facebook AI Research FAIR Proteins (Generative Biology) Team
Research Fellow Summer–Fall 2018	OpenAI Generative models

Awards and Honors

DARPA Young Faculty Award	2026
Outstanding Paper Finalist at RSS	2026
Outstanding Paper Award at ICLR	2024
National Science Foundation (NSF) Graduate Research Fellowship	2020-2024

Publications

The * symbol denotes equal contribution, and the † symbol denotes equal advising.

Conference Papers

- C131** Haoyu Chen, Kaichen Zhou, Hang Hua, Kaile Zhang, Jingwen Qian, Wufei Ma, Haonan Chen, Chunjiang Liu, Yizhou Zhao, Xiaoyuan Wang, Weiyue Li, Alan Yuille, Paul Pu Liang, Yilun Du. “Memo-Bench: Benchmarking World Modeling in Dynamically Changing Environments” *European Conference on Computer Vision (ECCV)*, 2026.
- C130** Kaichen Zhou*, Yuzhen Chen*, Fangneng Zhan, Hang Hua, Grace Chen, Xinhai Chang, Ao Qu, Yilun Du, Zhuang Liu, Paul Pu Liang†, Mengyu Wang†. “GEM-4D: Geometry-Enhanced Video World Models for Robot Manipulation” *European Conference on Computer Vision (ECCV)*, 2026.
- C129** Chaoqi Liu, Xiaoshen Han, Jiawei Gao, Yue Zhao, Haonan Chen, Yilun Du. “OAT: Ordered Action Tokenization” *Robotics: Science and Systems (RSS)*, 2026.

- C128** Zhiyi Li, Peilin Wu*, Xiaoshen Han*, Ruojin Cai, Yilun Du. “Structured 4D Latent Predictive Model for Robot Planning” *International Conference on Machine Learning (ICML)*, 2026.
- C127** Shuxin Cao, Liquan Wang, Walker Byrnes, Yiye Chen, Yilun Du, Animesh Garg. “Hierarchical Policy Learning via Spectral Decomposition” *International Conference on Machine Learning (ICML)*, 2026.
- C126** Haldun Balim, Na Li[†], Yilun Du[†]. “Model-Based Diffusion Sampling for Predictive Control in Offline Decision Making” *International Conference on Machine Learning (ICML)*, 2026.
- C125** Hansen Jin Lillemark*, Benhao Huang*, Fangneng Zhan, Yilun Du, T. Anderson Keller. “Flow Equivariant World Models: Memory for Partially Observed Dynamic Environments” *International Conference on Machine Learning (ICML)*, 2026.
- C124** Evgenii Opryshko*, Junwei Quan*, Claas Voelcker, Yilun Du, Igor Gilitschenski. “Test-Time Graph Search for Goal-Conditioned Reinforcement Learning” *International Conference on Machine Learning (ICML)*, 2026.
- C123** Renhao Zhang, Haotian Fu, Mingxi Jia, George Konidaris, Yilun Du, Bruno Castro da Silva. “From Noise to Control: Parameterized Diffusion Policies” *International Conference on Machine Learning (ICML)*, 2026.
- C122** Maijunxian Wang, Ruisi Wang, Juyi Lin, Ran Ji, Thaddäus Wiedemer, Qingying Gao, Dezhi Luo, Yaoyao Qian, Lianyu Huang, Zelong Hong, Jiahui Ge, Qianli Ma, Hang He, Yifan Zhou, Lingzi Guo, Lantao Mei, Jiachen Li, Hanwen Xing, Tianqi Zhao, Fengyuan Yu, Weihang Xiao, Yizheng Jiao, Jianheng Hou, Danyang Zhang, Pengcheng Xu, Boyang Zhong, Zehong Zhao, Gaoyun Fang, John Kitaoka, Yile Xu, Hua Xu, Kenton Blacutt, Tin Nguyen, Siyuan Song, Haoran Sun, Shaoyue Wen, Linyang He, Runming Wang, Yanzhi Wang, Mengyue Yang, Ziqiao Ma, Raphaël Milliére, Freda Shi, Nuno Vasconcelos, Daniel Khoshnab, Alan Yuille, Yilun Du, Ziming Liu, Bo Li, Dahua Lin, Ziwei Liu, Vikash Kumar, Yijiang Li, Lei Yang, Zhongang Cai, Hokin Deng. “A Very Big Video Reasoning Suite” *International Conference on Machine Learning (ICML)*, 2026.
- C121** Minda Zhao, Yilun Du, Mengyu Wang. “Large Language Models Are Bad Dice Players: LLMs Struggle to Generate Random Numbers from Statistical Distributions” *Annual Meeting of the Association for Computational Linguistics (ACL)*, 2026.
- C120** Zhiqiu Lin, Chancharik Mitra, Siyuan Cen, Isaac Li, Yuhan Huang, Yu Tong Tiffany Ling, Hewei Wang, Irene Pi, Shihang Zhu, Ryan Rao, George Liu, Jiaxi Li, Ruojin Li, Yili Han, Yilun Du, Deva Ramanan. “CHAI: Building a Precise Video Language with Human-AI Oversight” *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2026.
- C119** Haowen Liu*, Shaoxiong Yao*, Haonan Chen, Jiawei Gao, Jiayuan Mao, Jia-Bin Huang, Yilun Du. “SIMPACT: Simulation-Enabled Action Planning using Vision-Language Models” *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2026.
- C118** Yifan Liu, Fangneng Zhan, Kaichen Zhou, Yilun Du, Paul Pu Liang, Hanspeter Pfister. “Abstract 3D Perception for Spatial Intelligence in Vision-Language Models” *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2026.
- C117** Runhan Huang, Haldun Balim, Heng Yang, Yilun Du. “Flexible Locomotion Learning with Diffusion Model Predictive Control” *International Conference on Robotics and Automation (ICRA)*, 2026.
- C116** Haonan Chen, Jiaming Xu, Hongyu Chen, Kaiwen Hong, Binghao Huang, Chaoqi Liu, Jiayuan Mao, Yunzhu Li, Yilun Du[†], Katherine Driggs-Campbell[†]. “Multi-Modal Manipulation via Multi-Modal Policy Consensus” *International Conference on Robotics and Automation (ICRA)*, 2026.
- C115** Aayush Karan, Yilun Du. “Reasoning with Sampling: Your Base Model is Smarter Than You Think” *International Conference on Learning Representations (ICLR)*, 2026.

- C114** Chenyu Wang*, Zishen Wan*, Hao Kang, Emma Chen, Zhiqiang Xie, Tushar Krishna, Vijay Reddi, Yilun Du. “SLM-MUX: Orchestrating Small Language Models for Reasoning” *International Conference on Learning Representations (ICLR)*, 2026.
- C113** Xiangcheng Zhang, Haowei Lin, Haotian Ye, James Zou, Jianzhu Ma, Yitao Liang, Yilun Du. “Inference-time Scaling of Diffusion Models through Classical Search” *International Conference on Learning Representations (ICLR)*, 2026.
- C112** Jen-Yuan Huang, Tong Ling, Yilun Du. “Long-Text-to-Image Generation via Compositional Prompt Decomposition” *International Conference on Learning Representations (ICLR)*, 2026.
- C111** Jiahao Zhang*, Muqing Jiang*, Nanru Dai, Taiming Lu, Arda Uzunoglu, Shunchi Zhang, Yana Wei, Jiahao Wang, Vishal M. Patel, Paul Pu Liang, Daniel Khashabi, Cheng Peng, Rama Chellappa, Tianmin Shu, Alan Yuille, Yilun Du, Jieneng Chen. “World-in-World: World Models in a Closed-Loop World” *International Conference on Learning Representations (ICLR)*, 2026.
- C110** Yunzhi Zhang, Carson Murtuza-Lanier, Zizhang Li, Yilun Du, Jiajun Wu. “Product of Experts for Visual Generation” *International Conference on Learning Representations (ICLR)*, 2026.
- C109** Calvin Luo*, Zilai Zeng*, Mingxi Jia, Yilun Du, Chen Sun. “Self-Improving Loops for Visual Robotic Planning” *International Conference on Learning Representations (ICLR)*, 2026.
- C108** Yiming Li, Nael Darwiche, Amirreza Razmjoo, Sichao Liu, Yilun Du, Auke Ijspeert, Sylvain Calinon. “Geometry-aware Policy Imitation” *International Conference on Learning Representations (ICLR)*, 2026.
- C107** Alexi Gladstone, Ganesh Nanduru, Md Mofijul Islam, Peixuan Han, Hyeonjeong Ha, Aman Chadha, Yilun Du, Heng Ji, Jundong Li, Tariq Iqbal. “Energy-Based Transformers are Scalable Learners and Thinkers” *International Conference on Learning Representations (ICLR)*, 2026.
- C106** Jaeyeon Kim*, Lee Cheuk-Kit*, Carles Domingo-Enrich, Yilun Du, Sham Kakade, Timothy Ngotiaoco, Sitan Chen, Michael S. Albergo. “Any-Order Flexible Length Masked Diffusion” *International Conference on Learning Representations (ICLR)*, 2026.
- C105** Zhiqiu Lin*, Siyuan Cen*, Daniel Jiang, Jay Karhade, Hewei Wang, Chancharik Mitra, Yu Tong Tiffany Ling, Yuhan Huang, Sifan Liu, Mingyu Chen, Rushikesh Zawar, Xue Bai, Yilun Du, Chuang Gan, Deva Ramanan. “Towards Understanding Camera Motions in Any Video” *Neural Information Processing Systems (NeurIPS) Track on Datasets and Benchmarks*, 2025.
- C104** Zhenting Qi, Fan Nie, Alexandre Alahi, James Zou, Himabindu Lakkaraju, Yilun Du, Eric Xing, Sham Kakade, Hanlin Zhang. “EvoLM: In Search of Lost Language Model Training Dynamics” *Neural Information Processing Systems (NeurIPS)*, 2025.
- C103** Alexandru Oarga, Yilun Du. “Generalizable Reasoning through Compositional Energy Minimization” *Neural Information Processing Systems (NeurIPS)*, 2025.
- C102** Yunhao Luo, Utkarsh A. Mishra, Yilun Du[†], Danfei Xu[†]. “Generative Trajectory Stitching through Diffusion Composition” *Neural Information Processing Systems (NeurIPS)*, 2025.
- C101** Hao Kang, Qingru Zhang, Han Cai, Weiyuan Xu, Tushar Krishna, Yilun Du, Tsachy Weissman. “Win Fast or Lose Slow: Balancing Speed and Accuracy in Latency-Sensitive Decisions of LLMs” *Neural Information Processing Systems (NeurIPS)*, 2025.
- C100** Yuncong Yang*, Jiageng Liu*, Zheyuan Zhang, Siyuan Zhou, Reuben Tan, Jianwei Yang, Yilun Du, Chuang Gan. “MindJourney: Test-Time Scaling with World Models for Spatial Reasoning” *Neural Information Processing Systems (NeurIPS)*, 2025.
- C99** Siyuan Zhou, Yilun Du, Yuncong Yang, Lei Han, Peihao Chen, Dit-Yan Yeung, Chuang Gan. “Learning 3D Persistent Embodied World Models” *Neural Information Processing Systems (NeurIPS)*, 2025.

- C98** Louis Béthune, David Vigouroux, Yilun Du, Rufin VanRullen, Thomas Serre, Victor Boutin. “Follow the Energy, Find the Path: Riemannian Metrics from Energy-Based Models” *Neural Information Processing Systems (NeurIPS)*, 2025.
- C97** Yuncong Yang, Han Yang, Jiachen Zhou, Peihao Chen, Hongxin Zhang, Yilun Du, Chuang Gan. “3D-Mem: 3D Scene Memory for Embodied Exploration and Reasoning” *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025.
- C96** Shalev Lifshitz, Sheila A. McIlraith, Yilun Du. “Multi-Agent Verification: Scaling Test-Time Compute with Multiple Verifiers” *Conference on Language Modeling (CoLM)*, 2025.
- C95** Haoyu Zhen*, Qiao Sun*, Hongxin Zhang, Junyan Li, Siyuan Zhou, Yilun Du, Chuang Gan. “Tesser-Act: Learning 4D Embodied World Models” *International Conference on Computer Vision (ICCV)*, 2025.
- C94** Runhan Huang*, Shaoting Zhu*, Yilun Du, Hang Zhao. “MoELoco: Mixture of Experts for Multitask Locomotion” *International Conference on Intelligent Robots and Systems (IROS)*, 2025.
- C93** Yanbo Wang, Justin Dauwels, Yilun Du. “Compositional Scene Understanding through Inverse Generative Modeling” *International Conference on Machine Learning (ICML)*, 2025.
- C92** Kiwhan Song*, Boyuan Chen*, Max Simchowitz, Yilun Du, Russ Tedrake, Vincent Sitzmann. “History-Guided Video Diffusion” *International Conference on Machine Learning (ICML)*, 2025.
- C91** Shenyan Gao, Siyuan Zhou, Yilun Du, Jun Zhang, Chuang Gan. “Learning Adaptable World Models with Latent Actions” *International Conference on Machine Learning (ICML)*, 2025.
- C90** Zishen Wan, Jiayi Qian, Yuhang Du, Jason Jabbour, Yilun Du, Yang (Katie) Zhao, Arijit Raychowdhury, Tushar Krishna, Vijay Reddi. “Generative AI in Embodied Systems: System-Level Analysis of Performance, Efficiency and Scalability” *IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)*, 2025.
- C89** Sigmund H. Høeg, Yilun Du, Olav Egeland. “Streaming Diffusion Policy: Fast Policy Synthesis with Variable Noise Diffusion Models” *International Conference on Robotics and Automation (ICRA)*, 2025.
- C88** Yanwei Wang, Lirui Wang, Yilun Du, Balakumar Sundaralingam, Xuning Yang, Yu-Wei Chao, Claudia Perez-D’Arpino, Dieter Fox, Julie Shah. “Inference-Time Policy Steering through Human Interactions” *International Conference on Robotics and Automation (ICRA)*, 2025.
- C87** Siddharth Ancha, Sunshine Jiang, Travis Manderson, Laura Brandt, Yilun Du, Philip R. Osteen, Nicholas Roy. “Anomalies by Synthesis: Anomaly Detection using Generative Diffusion Models for Off-Road Navigation” *International Conference on Robotics and Automation (ICRA)*, 2025.
- C86** Vighnesh Subramaniam*, Yilun Du*, Joshua B. Tenenbaum, Antonio Torralba, Shuang Li, Igor Mordatch. “Multiagent Finetuning: Self Improvement with Diverse Reasoning Chains” *International Conference on Learning Representations (ICLR)*, 2025.
- C85** Yunhao Luo, Yilun Du. “Grounding Video Models to Actions through Goal Conditioned Exploration” *International Conference on Learning Representations (ICLR)*, 2025.
- C84** Calvin Luo, Zilai Zeng, Yilun Du, Chen Sun. “Solving New Tasks by Adapting Internet Video Knowledge” *International Conference on Learning Representations (ICLR)*, 2025.
- C83** Ying Fan, Yilun Du, Kannan Ramchandran, Kangwook Lee. “Looped Transformers for Length Generalization” *International Conference on Learning Representations (ICLR)*, 2025.
- C82** Wenhao Chai*, Enxin Song*, Yilun Du, Chenlin Meng, Vashisht Madhavan, Omer Bar-Tel, Jenq-Neng Hwang, Saining Xie, Christopher D. Manning. “AuroraCap: Efficient, Performant Video Detailed Captioning and a New Benchmark” *International Conference on Learning Representations (ICLR)*, 2025.

- C81** Hongxin Zhang*, Zeyuan Wang*, Qiushi Lyu*, Zheyuan Zhang, Sunli Chen, Tianmin Shu, Yilun Du, Chuang Gan. “COMBO: Compositional World Models for Embodied Multi-Agent Cooperation” *International Conference on Learning Representations (ICLR)*, 2025.
- C80** Julia Briden, Yilun Du, Enrico Zucchelli, Richard Linares. “Compositional Diffusion Models for Powered Descent Trajectory Generation with Flexible Constraints” *IEEE Aerospace Conference*, 2025.
- C79** Aviv Netanyahu, Yilun Du, Antonia Bronars, Jyotish Pari, Joshua B. Tenenbaum, Tianmin Shu, Pulkit Agrawal. “Few-Shot Task Learning through Inverse Generative Modeling” *Advances in Neural Information Processing Systems (NeurIPS)*, 2024.
- C78** Boyuan Chen, Diego Marti Monso, Yilun Du, Max Simchowitz, Russ Tedrake, Vincent Sitzmann. “Diffusion Forcing: Next-token Prediction Meets Full-Sequence Diffusion” *Advances in Neural Information Processing Systems (NeurIPS)*, 2024.
- C77** Yinan Zhang, Eric Tzeng, Yilun Du, Dmitry Kislyuk. “Large-scale Reinforcement Learning for Diffusion Models” *European Conference on Computer Vision (ECCV)*, 2024.
- C76** Jie Yin, Andrew Luo, Yilun Du, Anoop Cherman, Tim K. Marks, Jonathan Le Roux, Chuang Gan. “Disentangled Acoustic Fields For Multimodal Physical Scene Understanding” *International Conference on Intelligent Robots and Systems (IROS)*, 2024.
- C75** Yiqing Xu, Jiayuan Mao, Yilun Du, Tomás Lozano-Pérez, Leslie Kaelbling, David Hsu. ““Set It Up!”: Functional Object Arrangement with Compositional Generative Models” *Robotics: Science and Systems (RSS)*, 2024.
- C74** Lirui Wang, Alan Zhao, Yilun Du, Edward Adelson, Russ Tedrake. “PoCo: Policy Composition from and for Heterogeneous Robot Learning” *Robotics: Science and Systems (RSS)*, 2024.
- C73** Yilun Du, Leslie Kaelbling. “Compositional Generative Modeling: A Single Model is Not All You Need” *International Conference on Machine Learning (ICML)*, 2024.
- C72** Yilun Du, Shuang Li, Antonio Torralba, Joshua B. Tenenbaum, Igor Mordatch. “Improving Factuality and Reasoning in Language Models through Multiagent Debate” *International Conference on Machine Learning (ICML)*, 2024.
- C71** Yilun Du*, Jiayuan Mao*, Joshua B. Tenenbaum. “Learning Iterative Reasoning through Energy Diffusion” *International Conference on Machine Learning (ICML)*, 2024.
- C70** Yunhao Luo, Chen Sun, Joshua B. Tenenbaum, Yilun Du. “Potential Based Diffusion Motion Planning” *International Conference on Machine Learning (ICML)*, 2024.
- C69** Jocelin Su*, Nan Liu*, Yanbo Wang*, Joshua B. Tenenbaum, Yilun Du. “Compositional Image Decomposition with Diffusion Models” *International Conference on Machine Learning (ICML)*, 2024.
- C68** Siyuan Zhou, Yilun Du, Jiaben Chen, Yandong Li, Dit-Yan Yeung, Chuang Gan. “RoboDreamer: Learning Compositional World Models for Robot Imagination” *International Conference on Machine Learning (ICML)*, 2024.
- C67** Sherry Yang, Jacob Walker, Jack Parker-Holder, Yilun Du, Jake Bruce, Andre Barreto, Pieter Abbeel, Dale Schuurmans. “Video as the New Language for Real-World Decision Making” *International Conference on Machine Learning (ICML)*, 2024.
- C66** Haoyu Zhen, Xiaowen Qiu, Peihao Chen, Jincheng Yang, Xin Yan, Yilun Du, Yining Hong, Chuang Gan. “3D-VLA: A 3D Vision-Language-Action Generative World Model” *International Conference on Machine Learning (ICML)*, 2024.
- C65** Yilun Du, Sherry Yang, Pete Florence, Fei Xia, Ayzaan Wahid, Brian Ichter, Pierre Sermanet, Tianhe Yu, Pieter Abbeel, Joshua B. Tenenbaum, Leslie Kaelbling, Andy Zeng, Jonathan Tompson. “Video Language Planning” *International Conference on Learning Representations (ICLR)*, 2024.

- C64** Kevin Black*, Michael Janner*, Yilun Du, Ilya Kostrikov, Sergey Levine. “Training Diffusion Models with Reinforcement Learning” *International Conference on Learning Representations (ICLR)*, 2024.
- C63** Tailin Wu*, Takashi Maruyama*, Long Wei*, Tao Zhang*, Yilun Du*, Gianluca Iaccarino, Jure Leskovec. “Compositional Generative Inverse Design” *International Conference on Learning Representations (ICLR)*, 2024.
- C62** Mengjiao Yang*, Yilun Du*, Bo Dai, Dale Schuurmans, Joshua B. Tenenbaum, Pieter Abbeel. “Probabilistic Adaptation of Text-to-Video Models” *International Conference on Learning Representations (ICLR)*, 2024.
- C61** Mengjiao Yang, Yilun Du, Kamyar Ghasemipour, Jonathan Tompson, Leslie Kaelbling, Dale Schuurmans, Pieter Abbeel. “Learning Interactive Real-World Simulators” *International Conference on Learning Representations (ICLR)*, 2024.
- C60** Po-Chen Ko, Jiayuan Mao, Yilun Du, Shao-Hua Sun, Joshua B. Tenenbaum. “Learning to Act from Actionless Video through Dense Correspondences” *International Conference on Learning Representations (ICLR)*, 2024.
- C59** Yichen Li, Yilun Du, Chao Liu, Francis Williams, Michael Foshey, Benjamin Eckart, Jan Kautz, Joshua B. Tenenbaum, Antonio Torralba, Wojciech Matusik. “Learning to Jointly Understand Visual and Tactile Signals” *International Conference on Learning Representations (ICLR)*, 2024.
- C58** Hongxin Zhang, Weihua Du, Jiaming Shan, Qinhong Zhou, Yilun Du, Joshua B. Tenenbaum, Tianmin Shu, Chuang Gan. “Building Cooperative Embodied Agents Modularly with Large Language Models” *International Conference on Learning Representations (ICLR)*, 2024.
- C57** Qinhong Zhou, Sunli Chen, Yisong Wang, Haozhe Xu, Weihua Du, Hongxin Zhang, Yilun Du, Joshua B. Tenenbaum, Chuang Gan. “HAZARD Challenge: Embodied Decision Making in Dynamically Changing Environments” *International Conference on Learning Representations (ICLR)*, 2024.
- C56** Anurag Ajay*, Seungwook Han*, Yilun Du*, Shuang Li, Abhi Gupta, Tommi Jaakkola, Joshua B. Tenenbaum, Leslie Kaelbling, Akash Srivastava, Pulkit Agrawal. “Compositional Foundation Models for Hierarchical Planning” *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.
- C55** Yilun Du*, Mengjiao Yang*, Bo Dai, Hanjun Dai, Ofir Nachum, Joshua B. Tenenbaum, Dale Schuurmans, Pieter Abbeel. “Learning Universal Policies via Text-Guided Video Generation” *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.
- C54** Siyuan Zhou, Yilun Du, Shun Zhang, Mengdi Xu, Yikang Shen, Wei Xiao, Dit-Yan Yeung, Chuang Gan. “Adaptive Online Replanning with Diffusion Models” *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.
- C53** Cameron Smith, Yilun Du, Ayush Tewari, Vincent Sitzmann. “FlowCam: Training Generalizable 3D Radiance Fields without Camera Poses via Pixel-Aligned Scene Flow” *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.
- C52** Yining Hong, Haoyu Zhen, Peihao Chen, Shuhong Zheng, Yilun Du, Zhenfang Chen, Chuang Gan. “3D-LLM: Injecting the 3D World into Large Language Models” *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.
- C51** Tsun-Hsuan Wang, Juntian Zheng, Pingchuan Ma, Yilun Du, Byungchul Kim, Andrew Spielberg, Joshua B. Tenenbaum, Chuang Gan, Daniela Rus. “DiffuseBot: Breeding Soft Robots With Physics-Augmented Generative Diffusion Models” *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.
- C50** Shengzhuang Chen, Long-Kai Huang, Jonathan Richard Schwarz, Yilun Du, Ying Wei. “Secure Out-of-Distribution Task Generalization with Energy-Based Models” *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.

- C49 Zhutian Yang, Jiayuan Mao, Yilun Du, Jiajun Wu, Joshua B. Tenenbaum, Tomás Lozano-Pérez, Leslie Kaelbling. “Compositional Diffusion-Based Continuous Constraint Solvers” *Conference on Robot Learning (CoRL)*, 2023.
- C48 Nan Liu*, Yilun Du*, Shuang Li*, Joshua B. Tenenbaum, Antonio Torralba. “Unsupervised Compositional Concepts Discovery with Text-to-Image Generative Models” *International Conference on Computer Vision (ICCV)*, 2023.
- C47 Yilun Du, Conor Durkan, Robin Strudel, Joshua B. Tenenbaum, Sander Dieleman, Rob Fergus, Jascha Sohl-Dickstein, Arnaud Doucet, Will Grathwohl. “Reduce, Reuse, Recycle: Compositional Generation with Energy-Based Diffusion Models and MCMC” *International Conference on Machine Learning (ICML)*, 2023.
- C46 Armand Comas, Yilun Du, Christian Fernandez, Sandesh Ghimire, Mario Sznaiier, Joshua B. Tenenbaum, Octavia Camps. “Inferring Relational Potentials in Interacting Systems” *International Conference on Machine Learning (ICML)*, 2023.
- C45 Jiahui Fu, Yilun Du, Kurran Singh, Joshua B. Tenenbaum, John J. Leonard. “NeuSE: Neural SE(3)-Equivariant Embedding for Consistent Spatial Understanding with Objects” *Robotics: Science and Systems (RSS)*, 2023.
- C44 Weiyu Liu, Yilun Du, Tucker Hermans, Sonia Chernova, Chris Paxton. “StructDiffusion: Language-Guided Creation of Physically-Valid Structures using Unseen Objects” *Robotics: Science and Systems (RSS)*, 2023.
- C43 Cheng Chi, Siyuan Feng, Yilun Du, Zhenjia Xu, Eric Cousineau, Benjamin Burchfiel, Shuran Song. “Diffusion Policy: Visuomotor Policy Learning via Action Diffusion” *Robotics: Science and Systems (RSS)*, 2023.
- C42 Yilun Du, Cameron Smith, Ayush Tewari, Vincent Sitzmann. “Learning to Render Novel Views from Wide-Baseline Stereo Pairs” *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.
- C41 Yining Hong, Chunru Lin, Yilun Du, Zhenfang Chen, Joshua B. Tenenbaum, Chuang Gan. “3D Concept Learning and Reasoning from Multi-View Images” *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.
- C40 Anurag Ajay*, Yilun Du*, Abhi Gupta*, Joshua B. Tenenbaum, Tommi S. Jaakkola, Pulkit Agrawal. “Is Conditional Generative Modeling all You Need for Decision-Making?” *International Conference on Learning Representations (ICLR)*, 2023.
- C39 Shuang Li*, Yilun Du*, Joshua B. Tenenbaum, Antonio Torralba, Igor Mordatch. “Composing Ensembles of Pre-trained Models via Iterative Consensus” *International Conference on Learning Representations (ICLR)*, 2023.
- C38 Hongyi Chen*, Yilun Du*, Yiye Chen*, Joshua B. Tenenbaum, Patricio Antonio Vela. “Planning with Sequence Models through Iterative Energy Minimization” *International Conference on Learning Representations (ICLR)*, 2023.
- C37 Prafull Sharma, Ayush Tewari, Yilun Du, Sergey Zakharov, Rares Ambrus, Adrien Gaidon, William T. Freeman, Fredo Durand, Joshua B. Tenenbaum, Vincent Sitzmann. “Neural Groundplans: Persistent Neural Scene Representations from a Single Image” *International Conference on Learning Representations (ICLR)*, 2023.
- C36 Ethan Chun, Yilun Du, Anthony Simeonov, Tomás Lozano-Pérez, Leslie Kaelbling. “Local Neural Descriptor Fields: Locally Conditioned Object Representations for Manipulation” *International Conference on Robotics and Automation (ICRA)*, 2023.
- C35 Jose Iturralde*, Aiden Curtis*, Yilun Du, Leslie Kaelbling, Tomás Lozano-Pérez. “Visibility-Aware Navigation Among Movable Objects” *International Conference on Robotics and Automation (ICRA)*, 2023.

- C34** Anthony Simeonov*, Yilun Du*, Yen-Chen Lin, Alberto Rodriguez, Leslie Kaelbling, Tomás Lozano-Pérez, Pulkit Agrawal. “SE(3)-Equivariant Relational Rearrangement with Neural Descriptor Fields” *Conference on Robot Learning (CoRL)*, 2022.
- C33** Yen-Chen Lin, Pete Florence, Andy Zeng, Jonathan T. Barron, Yilun Du, Wei-Chiu Ma, Anthony Simeonov, Alberto Rodriguez, Phillip Isola. “MIRA: Mental Imagery for Robotic Affordances” *Conference on Robot Learning (CoRL)*, 2022.
- C32** Andrew Luo, Yilun Du, Michael J. Tarr, Joshua B. Tenenbaum, Antonio Torralba, Chuang Gan. “Learning Neural Acoustic Fields” *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.
- C31** Yining Hong, Yilun Du, Chunru Lin, Joshua B. Tenenbaum, Chuang Gan. “3D Concept Grounding on Neural Fields” *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.
- C30** Shuang Li, Xavier Puig, Chris Paxton, Yilun Du, Clinton Wang, Linxi Fan, Tao Chen, De-An Huang, Ekin Akyurek, Anima Anandkumar, Jacob Andreas, Igor Mordatch, Antonio Torralba, Yuke Zhu. “Pre-Trained Language Models for Interactive Decision-Making” *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.
- C29** Nan Liu*, Shuang Li*, Yilun Du*, Antonio Torralba, Joshua B. Tenenbaum. “Compositional Visual Generation with Composable Diffusion Models” *European Conference on Computer Vision (ECCV)*, 2022.
- C28** Yilun Du, Tomás Lozano-Pérez, Leslie Pack Kaelbling. “Learning Object-Based State Estimators for Household Robots” *International Conference on Intelligent Robots and Systems (IROS)*, 2022.
- C27** Jiahui Fu, Yilun Du, Kurran Singh, Joshua B. Tenenbaum, John J. Leonard. “Robust Change Detection Based on Neural Descriptor Fields” *International Conference on Intelligent Robots and Systems (IROS)*, 2022.
- C26** Michael Janner*, Yilun Du*, Joshua B. Tenenbaum, Sergey Levine. “Planning with Diffusion for Flexible Behavior Synthesis” *International Conference on Machine Learning (ICML)*, 2022.
- C25** Yilun Du, Shuang Li, Joshua B. Tenenbaum, Igor Mordatch. “Learning Iterative Reasoning through Energy Minimization” *International Conference on Machine Learning (ICML)*, 2022.
- C24** Rylan Schaeffer, Yilun Du, Gabrielle Liu, Ila Fiete. “Streaming Inference for Infinite Feature Models” *International Conference on Machine Learning (ICML)*, 2022.
- C23** Shuang Li, Yilun Du, Guido M. van de Ven, Antonio Torralba, Igor Mordatch. “Energy-Based Models for Continual Learning” *Conference on Lifelong Learning Agents (CoLLA)*, 2022.
- C22** Rylan Schaeffer, Gabrielle Liu, Yilun Du, Scott Linderman, Ila Fiete. “Streaming Inference for Infinite Non-Stationary Clustering” *Conference on Lifelong Learning Agents (CoLLA)*, 2022.
- C21** Klaus Greff, Francois Belletti, Lucas Beyer, Carl Doersch, Yilun Du, Daniel Duckworth, David J. Fleet, Dan Gnanapragasam, Florian Golemo, Charles Herrmann, Thomas Kipf, Abhijit Kundu, Dmitry Lagun, Issam Laradji, Derek Liu, Hsueh-Ti Liu, Henning Meyer, Yishu Miao, Derek Nowrouzezahrai, Cengiz Oztireli, Etienne Pot, Noha Radwan, Daniel Rebain, Sara Sabour, Mehdi Sajjadi, Matan Sela, Vincent Sitzmann, Austin Stone, Deqing Sun, Suhani Vora, Ziyu Wang, Tianhao Wu, Kwang Moo Yi, Fangcheng Zhong, Andrea Tagliasacchi. “Kubric: A Scalable Dataset Generator” *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.
- C20** Anthony Simeonov*, Yilun Du*, Andrea Tagliasacchi, Joshua B. Tenenbaum, Alberto Rodriguez, Pulkit Agrawal, Vincent Sitzmann. “Neural Descriptor Fields: SE(3)-Equivariant Object Representations for Manipulation” *International Conference on Robotics and Automation (ICRA)*, 2022.

- C19 Yilun Du, Shuang Li, Yash Sharma, Joshua B. Tenenbaum, Igor Mordatch. “Unsupervised Learning of Compositional Energy Concepts” *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.
- C18 Nan Liu*, Shuang Li*, Yilun Du*, Joshua B. Tenenbaum, Antonio Torralba. “Learning to Compose Visual Relations” *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.
- C17 Yilun Du, Katie Collins, Joshua B. Tenenbaum, Vincent Sitzmann. “Learning Signal-Agnostic Manifolds of Neural Fields” *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.
- C16 Joseph Suarez, Yilun Du, Clare Zhu, Igor Mordatch, Phillip Isola. “The Neural MMO Platform for Massively Multiagent Research” *Advances in Neural Information Processing Systems Track on Datasets and Benchmarks (NeurIPS)*, 2021.
- C15 Linqi Zhou, Yilun Du, Jiajun Wu. “3D Shape Generation and Completion through Point-Voxel Diffusion” *International Conference on Computer Vision (ICCV)*, 2021.
- C14 Yilun Du, Yinan Zhang, Hong-Xing Yu, Joshua B. Tenenbaum, Jiajun Wu. “Neural Radiance Flow for 4D View Synthesis and Video Processing” *International Conference on Computer Vision (ICCV)*, 2021.
- C13 Yilun Du, Chuang Gan, Phillip Isola. “Curious Representation Learning for Embodied Intelligence” *International Conference on Computer Vision (ICCV)*, 2021.
- C12 Shuang Li, Yilun Du, Antonio Torralba, Josef Sivic, Bryan Russell. “Weakly Supervised Human-Object Interaction Detection in Video via Contrastive Spatiotemporal Regions” *International Conference on Computer Vision (ICCV)*, 2021.
- C11 Yilun Du, Shuang Li, Joshua B. Tenenbaum, Igor Mordatch. “Improved Contrastive Divergence Training of Energy-Based Models” *International Conference on Machine Learning (ICML)*, 2021.
- C10 Yilun Du, Kevin Smith, Tomer Ullman, Joshua B. Tenenbaum, Jiajun Wu. “Unsupervised Discovery of 3D Physical Objects from Video” *International Conference on Learning Representations (ICLR)*, 2021.
- C9 Yilun Du, Shuang Li, Igor Mordatch. “Compositional Visual Generation with Energy Based Models” *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.
- C8 Anthony Simeonov, Yilun Du, Beomjoon Kim, Francois Hogan, Joshua B. Tenenbaum, Pulkit Agrawal, Alberto Rodriguez. “A Long Horizon Planning Framework for Manipulating Rigid Pointcloud Objects” *Conference on Robot Learning (CoRL)*, 2020.
- C7 Yilun Du, Joshua Meier, Jerry Ma, Rob Fergus, Alexander Rives. “Energy-based models for atomic-resolution protein conformations” *International Conference on Learning Representations (ICLR)*, 2020.
- C6 Xingyou Song, Yiding Jiang, Yilun Du, Behnam Neyshabur. “Observational Overfitting in Reinforcement Learning” *International Conference on Learning Representations (ICLR)*, 2020.
- C5 Joseph Suarez, Yilun Du, Igor Mordatch, Phillip Isola. “Neural MMO v1.3: A Massively Multiagent Game Environment for Training and Evaluating Neural Networks” *International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2020 (extended abstract).
- C4 Yilun Du, Toru Lin, Igor Mordatch. “Model-Based Planning with Energy-Based Models” *Conference on Robot Learning (CoRL)*, 2019.
- C3 Yilun Du, Igor Mordatch. “Implicit Generation and Modeling with Energy Based Models” *Advances in Neural Information Processing Systems (NeurIPS)*, 2019.
- C2 Yilun Du, Karthik Narasimhan. “Task-Agnostic Dynamics Priors for Deep Reinforcement Learning” *International Conference on Machine Learning (ICML)*, 2019.

- C1 Yilun Du, Zhijian Liu, Hector Basevi, Ales Leonardis, William T. Freeman, Joshua B. Tenenbaum, Jiajun Wu. “Learning to Exploit Stability for 3D Scene Parsing” *Advances in Neural Information Processing Systems (NeurIPS)*, 2018.

Journal Articles

- J10 Bohan Hou*, Gen Li*, Jindou Jia*, Tuo An*, Xinying Guo*, Sicong Leng, Haoran Geng, Yanjie Ze, Tatsuya Harada, Philip Torr, Oier Mees, Marc Pollefeys, Zhuang Liu, Jiajun Wu, Pieter Abbeel, Jitendra Malik, Yilun Du, Jianfei Yang. “World Model for Robot Learning: A Comprehensive Survey” *The International Journal of Robotics Research (IJRR)*, accepted, 2026.
- J9 Yubin Kim, Ken Gu, Chanwoo Park, Chunjong Park, Samuel Schmidgall, A. Ali Heydari, Yao Yan, Zhihan Zhang, Yuchen Zhuang, Yun Liu, Mark Malhotra, Paul Pu Liang, Hae Won Park, Yuzhe Yang, Xuhai Xu, Yilun Du, Shwetak Patel, Tim Althoff, Daniel McDuff, Xin Liu. “Capable language models can outgrow the benefits of collaboration” *Nature Machine Intelligence*, 2026.
- J8 Han Qi, Haocheng Yin, Aris Zhu, Yilun Du, Heng Yang. “Inference-Time Enhancement of Generative Robot Policies via Predictive World Modeling” *IEEE Robotics and Automation Letters (RA-L)*, 2026.
- J7 Chaoqi Liu, Haonan Chen, Sigmund H. Høeg, Shaoxiong Yao, Yunzhu Li, Kris Hauser, Yilun Du. “Flexible Multitask Learning With Factorized Diffusion Policy” *IEEE Robotics and Automation Letters (RA-L)*, 2026.
- J6 Sigmund H. Høeg, Aksel Vaaler, Chaoqi Liu, Olav Egeland, Yilun Du. “Hybrid Diffusion for Simultaneous Symbolic and Continuous Planning” *IEEE Robotics and Automation Letters (RA-L)*, 2026.
- J5 Jiahui Fu, Yilun Du, Kurran Singh, Joshua B. Tenenbaum, John J. Leonard. “NeuSE: Neural SE(3)-Equivariant Embedding for Long-Term Object-Based Simultaneous Localization and Mapping” *The International Journal of Robotics Research (IJRR)*, 2026. *Extended journal version of the RSS 2023 paper.*
- J4 Julen Urain, Ajay Mandlekar, Yilun Du, Nur Muhammad “Mahi” Shafiullah, Danfei Xu, Katerina Fragkiadaki, Georgia Chalvatzaki, Jan Peters. “A Survey on Deep Generative Models for Robot Learning From Multimodal Demonstrations” *IEEE Transactions on Robotics (T-RO)*, 2026.
- J3 Robert McCarthy, Daniel C. H. Tan, Dominik Schmidt, Fernando Acero, Nathan Herr, Yilun Du, Thomas G. Thuruthel, Zhibin Li. “Towards Generalist Robot Learning from Internet Video: A Survey” *Journal of Artificial Intelligence Research (JAIR)*, 2025.
- J2 Yiqing Xu, Jiayuan Mao, Linfeng Li, Yilun Du, Tomás Lozano-Pérez, Leslie Pack Kaelbling, David Hsu. “‘Set It Up’: Functional Object Arrangement with Compositional Generative Models” *The International Journal of Robotics Research (IJRR)*, 2025. *Extended journal version of the RSS 2024 paper.*
- J1 Cheng Chi*, Zhenjia Xu*, Siyuan Feng, Eric Cousineau, Yilun Du, Benjamin Burchfiel, Russ Tedrake, Shuran Song. “Diffusion Policy: Visuomotor Policy Learning via Action Diffusion” *The International Journal of Robotics Research (IJRR)*, 2025. *Extended journal version of the RSS 2023 paper.*

Selected Preprints

- A8 Xiangcheng Zhang, Yilun Du. “World Action Planner: Generalizable Decision-Making with Action-Conditioned World Models” *arXiv preprint arXiv:2607.27599*, 2026. [Paper](#).
- A7 Alexi Gladstone, Heng Ji, Yilun Du. “Explorative Modeling: Unlocking a Third Pretraining Axis and End-to-End Generation” *arXiv preprint arXiv:2607.27372*, 2026. [Paper](#).

- A6** Zhenting Qi, Huangyuan Su, Ao Qu, Chenyu Wang, Yu Yao, Han Zheng, Kushal Chattopadhyay, Guowei Xu, Zihan Wang, Weirui Ye, Vijay Janapa Reddi, Ju Li, Paul Pu Liang, Himabindu Lakkaraju, Sham Kakade, Yilun Du. “Economy of Minds: Emerging Multi-Agent Intelligence with Economic Interactions” *arXiv preprint arXiv:2606.02859*, 2026. [Paper](#).
- A5** Seonghyeon Ye, Yunhao Ge, Kaiyuan Zheng, Shenyuan Gao, Sihyun Yu, George Kurian, Suneel Indupuru, You Liang Tan, Chuning Zhu, Jiannan Xiang, Ayaan Malik, Kyungmin Lee, William Liang, Nadun Ranawaka, Jiasheng Gu, Yinzhen Xu, Guanzhi Wang, Fengyuan Hu, Avnish Narayan, Johan Bjorck, Jing Wang, Gwanghyun Kim, Dantong Niu, Ruijie Zheng, Yuqi Xie, Jimmy Wu, Qi Wang, Ryan Julian, Danfei Xu, Yilun Du, Yevgen Chebotar, Scott Reed, Jan Kautz, Yuke Zhu, Linxi “Jim” Fan, Joel Jang. “World Action Models are Zero-shot Policies” *arXiv preprint arXiv:2602.15922*, 2026. [Paper](#).
- A4** Mingyang Deng, He Li, Tianhong Li, Yilun Du, Kaiming He. “Generative Modeling via Drifting” *arXiv preprint arXiv:2602.04770*, 2026. [Paper](#).
- A3** Boyuan Chen*, Tianyuan Zhang*, Haoran Geng*, Kiwhan Song, Caiyi Zhang, Peihao Li, William T. Freeman, Jitendra Malik, Pieter Abbeel, Russ Tedrake, Vincent Sitzmann, Yilun Du. “Large Video Planner Enables Generalizable Robot Control” *arXiv preprint arXiv:2512.15840*, 2025. [Paper](#).
- A2** Runqian Wang, Yilun Du. “Equilibrium Matching: Generative Modeling with Implicit Energy-Based Models” *arXiv preprint arXiv:2510.02300*, 2025. [Paper](#).
- A1** Robert Verkuil*, Ori Kabeli*, Yilun Du, Basile Wicky, Lukas Milles, Justas Dauparas, David Baker, Sergey Ovchinnikov, Tom Sercu, Alexander Rives. “Language Models Generalize Beyond Natural Proteins” *bioRxiv preprint*, 2022. [Paper](#).

Technical Reports

- P6** Hirokatsu Kataoka, Yoshihiro Fukuhara, Yonglong Tian, Shangzhe Wu, Oishi Deb, Ryosuke Yamada, Christian Rupprecht, Jianyuan Wang, Kohsuke Ide, Koichi Namekata, Xianzheng Ma, Yiming Chen, Robert Geirhos, Aditi Raghunathan, Yuki M. Asano, Deva Ramanan, David Fouhey, Andrew J. Davison, Yilun Du, Jiajun Wu, Zhuang Liu. “Visual General Intelligence: A White Paper” *arXiv preprint arXiv:2608.25924*, 2026. [Paper](#).
- P5** NVIDIA. “Cosmos 3: Omnimodal World Models for Physical AI” *arXiv preprint arXiv:2606.02800*, 2026. [Website Paper](#).
- P4** Gemini Robotics Team. “Evaluating Gemini Robotics Policies in a Veo World Simulator” *arXiv preprint arXiv:2512.10675*, 2025. [Website Paper](#).
- P3** Gemini Robotics Team. “Gemini Robotics 1.5: Pushing the Frontier of Generalist Robots with Advanced Embodied Reasoning, Thinking, and Motion Transfer” *arXiv preprint arXiv:2510.03342*, 2025. [Blog Paper](#).
- P2** Gemini Team. “Gemini 2.5: Pushing the Frontier with Advanced Reasoning, Multimodality, Long Context, and Next Generation Agentic Capabilities” *arXiv preprint arXiv:2507.06261*, 2025. [Blog Paper](#).
- P1** Gemini Robotics Team. “Gemini Robotics: Bringing AI into the Physical World” *arXiv preprint arXiv:2503.20020*, 2025. [Blog Paper](#).

Theses

- T2** Yilun Du. “Learning Generalizable Systems by Learning Composible Energy Landscapes.” Ph.D. thesis, Massachusetts Institute of Technology, 2024.
- T1** Yilun Du. “Implicit Learning with Energy-Based Models.” M.S. thesis, Massachusetts Institute of Technology, 2020.

Teaching

COMPSCI 1810: Machine Learning (Undergraduate) Co-Instructor, Harvard University	Spring 2026
COMPSCI 2821R: Topics in Machine Learning: Compositional AI Systems (Graduate Seminar) Instructor, Harvard University	Fall 2025

Invited Talks

WM@Booth 2026, Chicago Booth Embodied Reasoning with World Models	Aug. 2026
2026 IAIFI Summer Workshop Building Intelligence through Energy Based Models	Aug. 2026
MSRNE Generative Modeling & Sampling Workshop Reasoning with Generative Models through Energy Minimization	Aug. 2026
Diffusion Generative Modeling: Progress and Next Steps, Simons Institute Compositional Reasoning with Diffusion Models	Aug. 2026
Video Model Journal Club Embodied Reasoning with World Models	Jul. 2026
RSS Workshop on Diffusion for Robot Learning Model-Based Decision-Making with Diffusion Models	Jul. 2026
RSS Workshop on Robot Planning in the Era of Foundation Models Planning with World Models	Jul. 2026
CVPR Workshop on Embodied Reasoning in Action (ERA) Embodied Reasoning with World Models	Jun. 2026
CVPR Embodied AI Workshop Robot Manipulation through World Models	Jun. 2026
4th CVPR Workshop on Generative Models for Computer Vision Visual Scene Understanding through Inverse Generative Modeling	Jun. 2026
Third CVPR Workshop on Visual Concepts Compositional Energy-Based Concepts	Jun. 2026
CVPR Workshop on GeoFreeNVS: Geometry-Free Novel View Synthesis and Controllable Video Models Unified Generative-Predictive Modeling for 4D Scene Understanding	Jun. 2026
CVPR Workshop on Emerging Directions in Data for Multimodal Foundation Models A Very Big Video Reasoning Suite	Jun. 2026
1st CVPR Workshop on Video World Models: Interaction, Memory, and Efficiency Robot Manipulation with World Models	Jun. 2026
CVPR Workshop on From Labs to Life: Embodied Intelligence in the Wild Embodied Intelligence with World Models	Jun. 2026

Brown University Workshop on Agentic Scientific Computing and Scientific Machine Learning Multiagent Scientific Reasoning Systems	May 2026
Oregon State University Building Embodied Intelligence with Compositional World Models	Jan. 2026
NeurIPS Workshop on Embodied World Models Building Intelligent Robots with World Models	Dec. 2025
NeurIPS Tutorial on Beyond Autoregressive Models Autoregressive World Models	Dec. 2025
MIT FutureTech AI Automation with Generative World Models	Nov. 2025
Amazon Robotics Robot Manipulation with World Models	Nov. 2025
Tufts ECE Seminar Learning Compositional Models of the World	Sep. 2025
NVIDIA Multiagent Reasoning with Language Models	Aug. 2025
Toyota Research Institute ML and Robotics Seminar Flexible Decision-Making with Compositional Constraints	Aug. 2025
Meta MRS AI Seminar Multiagent Reasoning with Language Models	Jul. 2025
Microsoft Research Seminar on Sampling and Generative Models Inference-Time Reasoning with Generative Models	Jul. 2025
ICML Workshop on Multiagent Systems Multiagent Language Model Reasoning Systems	Jul. 2025
RSS Workshop on Leveraging Foundation Models for Interactive Robot Learning Scaling Robot Learning with Video Models	Jun. 2025
CVPR Tutorial on Continual Robot Learning from Humans Continual Robot Learning with Compositional Generative Models	Jun. 2025
CVPR Tutorial on Robotics 101: An Odyssey from a Vision Perspective Learning Compositional Models of the World	Jun. 2025
CVPR Workshop on 3D-LLM/VLA Building 4D Models of the World	Jun. 2025
CVPR Workshop on Foundation Models Meet Embodied Agents Learning Compositional Models of the World	Jun. 2025
Kempner Institute Frontiers in NeuroAI Symposium Learning Compositional Models of the World	Jun. 2025
ICRA Workshop on Foundation Models and Neuro-Symbolic AI for Robotics Generalizable Foundation Models through Compositional Generation	May 2025
New England NLP Symposium Constructing Generalizable Multimodal Models through Compositional Generation	Apr. 2025
Weizmann AI and Vision Seminar Learning Compositional Models of the World	Mar. 2025

Shanghai AI Lab Learning Compositional Models of the World	Nov. 2024
Generalizing Outside the Training Distribution through Compositional Generation <i>2024:</i> UCLA (November); KAIST (October); Vanderbilt Machine Learning (September); NUS Tea Time (July); CMU MLD, University of Michigan CSE, UT Austin, UBC, Princeton, and Brown (April); NYU, Columbia, UPenn, University of Washington, Georgia Tech, University of Toronto, and Harvard (March); UCSD and Boston Dynamics AI Institute (February); ASU Frontiers in GenAI Seminar and Genesis Therapeutics (January). <i>2023:</i> JHU, KAIST, and CMU AI Seminar (November); Brown University Robotics Seminar, Cornell University Robotics Seminar, UC Berkeley, and UT Austin (October); Stanford Vision Seminar (July).	2023–2024
ICML Workshop on Multimodal Foundation Models in Embodied AI Compositional Foundation Models	Jul. 2024
RSS Workshop on Semantic Reasoning and Goal Understanding in Robotics Panelist	Jul. 2024
RSS Workshop on GenAI and HRI Building Customizable Generative Models through Compositional Generation	Jul. 2024
NeurIPS Tutorial on Language Models Meet World Models Panelist	Dec. 2023
BuzzRobot Learning a Visual Imagination	Dec. 2023
University of Tokyo, IRCN Keynote Speech Towards Multi-AI Intelligent Systems	Oct. 2023
Eye on AI Podcast AI Self-Improvement from Multiagent Debate	Oct. 2023
University of Montreal Robot Learning Seminar Riding the Wave of Generative Modeling for Decision-Making	Oct. 2023
Planning with Diffusion Models Donders Institute (February 2023); DeepMind (August 2022)	2022–2023
Implicit Learning with Energy-Based Models Nuro (October); Tencent (September); CMU (August)	2022
Generally Intelligent Podcast Modularity and Energy-Based Models	Nov. 2021

Advising

Co-advisers and joint mentors are listed in parentheses.

Postdoctoral Researchers Ruojin Cai (Kempner Research Fellow); Haonan Chen; Anand Gopalakrishnan (jointly advised with Samuel Gershman); Weirui Ye (MIT CSAIL; jointly advised with Leslie Pack Kaelbling and Tomás Lozano-Pérez)	Current
---	---------

Ph.D. Students

Primary advisees: Amani Kiruga; Hang Le; Runhan Huang; Zhiyi Li (MIT EECS).
Co-advised students: Zhenting Qi (with Hima Lakkaraju); Haldun Balim (with Na Li); Sarah Liaw (with David Alvarez-Melis); Aayush Karan (with Sitan Chen); Haoyu Xiong (MIT EECS; with Antonio Torralba); Chenyu Wang (Harvard ECE; with Vijay Janapa Reddi)

Current

Professional Service

Conference Area Chair

Robotics Science and Systems (RSS)	2026
Neural Information Processing Systems (NeurIPS)	2026
International Conference in Machine Learning (ICML)	2026
Association for the Advancement of Artificial Intelligence (AAAI)	2026
International Conference in Learning Representations (ICLR)	2025-26
Conference on Robot Learning (CoRL)	2025-26
Conference on Language Modeling (CoLM)	2025-26
International Conference in Computer Vision (ICCV)	2025

Conference Refereeing

Neural Information Processing Systems (NeurIPS)	2021-25
International Conference in Machine Learning (ICML)	2021-25
International Conference in Learning Representations (ICLR)	2021-24
Computer Vision and Pattern Recognition (CVPR)	2021-23
International Conference in Computer Vision (ICCV)	2021,23
European Conference in Computer Vision (ECCV)	2022
Special Interest Group in Graphics (SIGGRAPH)	2022-24
SIGGRAPH Asia	2024
Conference on Robot Learning (CoRL)	2023-24
Robotics: Science and Systems (RSS)	2023
International Conference on Robotics and Automation (ICRA)	2022-23
International Conference on Intelligent Robots and Systems (IROS)	2022-23

Workshop Organization

2nd Embodied Spatial Reasoning (ESR) Workshop (NeurIPS 2026)	Dec. 2026
Small Data, Rich Sensing: Multimodal Learning for Robotic Manipulation (CoRL 2026)	Nov. 2026
Remember, Retrieve, Act: Memory for Robot Foundation Models (CoRL 2026)	Nov. 2026
Compositional Generalization for Real-World Robot Learning (CoRL 2026)	Nov. 2026
Compositional and Modular Learning in the Era of Scaling in Robotics (IROS 2026)	Sep. 2026
Tactile Sensing for Robotic Foundation Models (RSS 2026)	Jul. 2026
2nd Workshop on Compositional Learning: Safety, Interpretability, and Agents (ICML 2026)	Jul. 2026
The 2nd 3D-LLM/VLA Workshop: Bridging Language, Vision and Action in 3D Environments (CVPR 2026)	Jun. 2026
Workshop on World Models Meet Active Sensing and Closed-Loop Planning (CVPR 2026)	Jun. 2026
1st Workshop on Multi-Agent Robotic Systems: Scaling with Compositional Intelligence (CVPR 2026)	Jun. 2026
Workshop on Robotics World Modeling (CoRL 2025)	Sep. 2025

Building Physically Plausible World Models (ICML 2025)	Jul. 2025
Robot Planning in the Era of Foundation Models (FM4RoboPlan) (RSS 2025)	Jun. 2025
Compositional Learning: Perspectives, Methods, and Paths Forward (NeurIPS 2024)	Dec. 2024
Safe Generative AI (NeurIPS 2024)	Dec. 2024
Geometric and Algebraic Structure in Robot Learning (RSS 2024)	Jul. 2024
Generative Models for Decision Making (ICLR 2024)	May 2024
Foundation Models for Decision Making (NeurIPS 2023)	Dec. 2023
Foundation Models for Decision Making (NeurIPS 2022)	Dec. 2022

Seminar Organization

MIT Embodied Intelligence Seminar	2021-2023
-----------------------------------	-----------